Global optimization on set of mixed variables: continuous and discrete with unordered possible values

***A S Mikhalev, **A I Rouban**

Computer Science Dept, Institute of Space and Information Technologies, Siberian Federal University, Street of academician Kirenskogo, nimber 26, Krasnoyarsk region, Krasnoyarsk 660 074, Russia

E-mail: * spaming1@yandex.ru; ** ai-rouban@mail.ru

Abstract. The algorithms of global non-differentiable minimization of functions on set of the mixed variables: continuous and discrete with unordered specific possible values are constructed. The method of optimization is based on selective averaging of required variables, on adaptive reorganization of the sizes of admissible domain of trial movements and on use of relative values for minimised functions. Existence of discrete variables leads to solution of a sequence of global minimization problems of the functions in space of only continuous variables at the presence: 1) of their inequality restrictions for each problem; 2) of the general inequality restrictions for all problems (i.e. at the absence of dependence of functions fore inequality restrictions from discrete variables). In the first case, presence of discrete variables with unordered non-numeric possible values leads to solution of sequence of problems of global minimization of multiextreme functions on set only of continuous variables at the presence of their inequality restrictions. As a result, among the received optimum solutions the best is selected. In the second variant all minimized functions is convoluted in each sampling point in one multiextreme function and this function is minimised on continuous variables.

1. Introduction

The problem of search of a global extremum [1-7] is very difficult. Its specificity are caused by possible breaks in many-extreme optimised functions and in functions of restrictions, the presence of noises, the limited set of possible values of the required variables and also the presence of discrete variables in addition to continuous variables.

The area of the mixed global optimization is applied for solution of optimization problems in different scopes. Among them are process industry and the financial engineering, engineering design, chemical engineering, the theory of control and process operations research, automobile and aircraft manufacturing and many others. Application of algorithms of mixed global optimization approach allows to solution problems of calculation of optimal parameters of the equipment, optimal conditions of functioning of process, optimal design of networks, building of schemes, etc.

Methods of solution of problems of search of optimal value in space continuous and discrete variables are discussed in the literature since the early 1980’s. Approaches to their solution can be mainly separated into two classes. The first class – is heuristic methods, which don’t guarantee that when the search is complete, the found solution will be optimal. Other class – deterministic methods – in turn, stop search with the guaranteed solution or indicate, that the problem has no integer solution.

The deterministic methods for solution of this class problem usually are based on the successive solutions of closely related problems of nonlinear programming. The basic concept which is underlies the algorithms of solution of problems – generation and refinement of estimates of optimal solution. Algorithms differ on way of forming of these estimates and on type of sequence of subtasks, which are solved for receiving these estimates [8-11].

Example of the heuristic approach to solution of problems is the method, which combined together a global search phase and a local search phase [12].

In this paper on the basis of a method [13–15] the algorithms of the decision of two problems of global optimisation with the mixed variables are received.

Discrete variables, based on properties of their possible values, conditionally can be divided into 2 classes: with **unordered** and **ordered** possible values. In turn, possible values of discrete variables of both classes can accept **non-numerical** and **numerical** possible values.

In this paper discrete variables with disorder non-numerical possible values are considered, and for each of them there is a minimized function and the inequality restrictions. Restrictions can be and the general for all minimized functions.

In the presence of discrete variables with the ordered possible values in other papers the approach based on one-for-one transition from discrete variable to a continuous variable will be considered. To each possible value of a discrete variable the continuous interval of values is put in conformity. Transition from continuous values to the discrete is carried out both at formation of trial movements and at calculation of required variables on each working step. This way allows to apply a method of global optimization in space of continuous variables to optimization on set mixed variables.

2. Statement of problem

The problem of search of a conditional global minimum of objective function of many variables $f(x,y)$ on set of the mixed variables: continuous and discrete with unordered specific possible values at the presence of inequality restrictions is solved:

$$f(x,y)=\min,\quad \varphi_j(x,y)\le 0,\ j=\overline{1,m(y)}. \tag{1}$$

Here: $x=(x_1,\dots,x_k)$ – vector k of continuous variables, y – discrete variable with r unordered non-numerical possible values $y_1,\dots,y_r$.

The presence of additional discrete variable with unordered possible values leads to solution of r independent problems of global minimization (1) only on continuous variables at the presence of inequality restrictions:

$$\left[f_\mu(x)=\min,\ \varphi_{j\mu}(x)\le 0,\ j=\overline{1,m_\mu}\right],\ \mu=\overline{1,r}. \tag{2}$$

Here:

$$f(x,y_\mu)\equiv f_\mu(x),\ \ \varphi_j(x,y_\mu)\equiv\varphi_{j\mu}(x).$$

As a result, among the obtained optimum solutions $x^*_\mu,\ \mu=\overline{1,r}$ the best solution $x^*_{\mu^*}$ is selected. It corresponds to the optimum value $y^*_{\mu^*}$ (with number μ^*) of a discrete variable.

3. The method selective averaging of required variables

We provide the solution of a r problems of global optimization (2) by using basic algorithm of method with selective averaging of required continuous variables [13-15]. The method is based on separation in time of trial and working steps, uniform distribution of sampling points in admissible region, selective averaging of required variables by results of the experimental data, obtained in sampling points, adaptive reorganization at each working step of the sizes of rectangular domain of trial movements. Further we will describe algorithm of finding of a global minimum (2) at the fixed μ.

For receipt of sampling points $x_{v,\mu}^{(i)},\ i=\overline{1,n}$ in rectangular domain Π_{μ}^{l} with a center in a point $x_{v,\mu}^{l}$ is sequentially generated uniform distribution of points:

$$x_{v,\mu}^{(i)} = x_{v,\mu}^{l} + \Delta x_{v,\mu}^{l} u_{v}^{(i)},\ u_{v} \in [-1;1],\ v=\overline{1,k},\ i=1,2,\ldots \quad (3)$$

and from them points, which satisfy to inequalities restrictions, are left. These points are in admissible region $X_{\mu}^{l} = \Pi_{\mu}^{l} \cap X_{\mu}$.

The starting point x_{μ}^{0} and sizes $\Delta x_{v,\mu}^{0}$ of rectangular domain Π_{μ}^{0} of trial movements are selected so, that Π_{μ}^{0} covered the admissible region X_{μ} or the part where the required global minimum is located.

In sampling points the minimized functions $f_{\mu}^{(i)} \equiv f_{\mu}(x^{(i)}),\ i=\overline{1,n}$ are calculated. After, the position of minimum is specified.

At movement to the extremum occurs adaptation (most of reduction) of the region of search movements. New value $x_{v,\mu}^{l+1}$, on average closer to position of a global minimum of the optimized function f_{μ}, and the sizes of rectangular domain of search movements $\Delta x_{v,\mu}^{l+1},\ v=\overline{1,k}$, are calculated on formulas [14]:

$$x_{v,\mu}^{l+1} = x_{v,\mu}^{l} + \Delta x_{v,\mu}^{l} u_{v,\mu,\min},\ u_{v,\mu,\min} = \sum_{i=1}^{n} u_{v,\mu}^{(i)} \overline{p}_{s,\min}^{(i)},\ v=\overline{1,k}, \quad (4)$$

$$\overline{p}_{s,\min}^{(i)} = \frac{p_s(g_{\mu,\min}^{(i)})}{\sum_{j=1}^{n} p_s(g_{\mu,\min}^{(j)})},\ g_{\mu,\min}^{(i)} = \frac{f_{\mu}^{(i)} - \widehat{f}_{\mu,\min}}{\widehat{f}_{\mu,\max} - \widehat{f}_{\mu,\min}},$$

$$\Delta x_{v,\mu}^{l+1} = \gamma_q \Delta x_{v,\mu}^{l} \Big(\sum_{i=1}^{n} |u_{v,\mu}^{(i)}|^q\ \overline{p}_{s,\min}^{(i)}\Big)^{1/q},\ v=\overline{1,k},$$

$$l=0,1,2,\ldots;[0<\gamma_q, q\in\{1,2,\ldots\}, 0<s].$$

Here $\widehat{f}_{\mu,\max} = \max\{f_{\mu}^{(i)}, i=\overline{1,n}\}$, $\widehat{f}_{\mu,\min} = \min\{f_{\mu}^{(i)}, i=\overline{1,n}\}$, $p_s(\cdot)$ – kernel, s – a degree of selectivity (ore selectivity coefficient) of kernel $p_s(\cdot)$. The kernels are normalized on 1 on a system of n sampling points: $\sum_{i=1}^{n} \overline{p}_{s,\min}^{(i)} = 1$. The argument of kernel – is non-dimensional variable, it always lies in interval $[0;1]$: $0 \le g_{\mu,\min}^{(i)} \le 1$.

When approaching to minimum the region of trial movement reduced, and thus, there is more exact tracking of position of extremum. The criterion of stop of search process is the condition of reduction (at some l) of size of region of variation of variables to the set value:

$$\max\left\{\frac{\Delta x_{v,\mu}^{l}}{\Delta x_{v,\mu}^{0}}, v=\overline{1,k}\right\} \le \varepsilon \ \text{or}\ \frac{\Delta x_{v,\mu}^{l}}{\Delta x_{v,\mu}^{0}} \le \varepsilon, v=\overline{1,k}. \quad (5)$$

After finding of global minimum for each of r functions $\{x_{\mu}^{*},\ f_{\mu}(x_{\mu}^{*}),\ \mu=\overline{1,r}\}$ the best solution $x_{\mu^*}^{*},\ f_{\mu^*}(x_{\mu^*}^{*})$ is selected:

$$\min\{f_{\mu}(x_{\mu}^{*}), \mu=\overline{1,r}\} = f_{\mu^*}(x_{\mu^*}^{*}). \quad (6)$$

Optimal value of discrete variable has number μ^*.

4. Example 1

We form 4 thirteen-extreme functions:

$\mu=1$: $f_1(x_1,x_2)=\min\{4|x_1-2|+5|x_2-2|+16;\ 6|x_1-2|^{1.1}+7|x_2-6|^{0.7}+4;$
$3|x_1-2|^{1.3}+3|x_2-10|^{1.3}+10;\ 7|x_1-6|^{0.8}+7|x_2-10|^{0.6}+2;$
$5|x_1-10|^{1.8}+5|x_2-10|^{1.8}+14;\ 8|x_1-10|^{0.9}+8|x_2-6|^{1.1}+6;$
$6|x_1-10|^{0.5}+7|x_2-2|^{0.5}+12;\ 2|x_1-6|^{1.8}+5|x_2-2|^{1.5}+8;$
$6|x_1-4|^{0.9}+5|x_2-4|^{0.7}+9;\ 4|x_1-4|^{1.2}+4|x_2-8|^{1.5}+3;$
$6|x_1-8|^{1.6}+5|x_2-8|^{0.6}+7;\ 5|x_1-8|^{1.3}+7|x_2-4|^{0.9}+5;\ 5|x_1-6|^{2}+5|x_2-6|^{2}\}$;
$\mu=2$: $f_2(x_1,x_2)=\min\{3|x_1+2|^{1.1}+3|x_2-2|^{1.3}+9;\ 7|x_1+2|^{0.9}+6|x_2-6|^{0.7}+4;$
$3|x_1+2|^{0.9}+4|x_2-10|^{0.3}+7;\ 6|x_1+6|^{0.9}+6|x_2-10|^{0.3}+9;$
$5|x_1+10|^{0.7}+3|x_2-10|^{1.6}+5;\ 3|x_1+10|^{0.8}+5|x_2-6|^{1.7}+6;$
$4|x_1+10|^{0.9}+2|x_2-2|^{1.1}+11;\ 2|x_1+6|^{1.8}+3|x_2-2|^{1.8}+8;$
$6|x_1+4|^{1.5}+5|x_2-4|^{1.5}+19;\ 4|x_1+4|^{1.3}+4|x_2-8|^{1.3}+13;$
$3|x_1+8|^{0.9}+4|x_2-8|^{0.9}+15;\ 4|x_1+8|^{0.6}+7|x_2-4|^{0.6}+17;\ 6|x_1+6|^{0.8}+6|x_2-6|^{0.6}+2\}$;
$\mu=3$: $f_3(x_1,x_2)=\min\{4|x_1+2|^{2.4}+5|x_2+2|^{1.6}+7;\ 3|x_1+2|^{1.8}+2|x_2+6|^{2.2}+14;$
$3|x_1+2|^{1.7}+3|x_2+10|^{1.9}+9;\ 4|x_1+6|^{1.4}+2|x_2+10|^{1.8}+10;$
$6|x_1+10|^{0.8}+7|x_2+10|^{1.5}+13;\ 5|x_1+10|^{1.3}+3|x_2+6|^{1.4}+8;$
$4|x_1+10|^{1.6}+5|x_2+2|^{1.3}+11;\ 2|x_1+6|^{1.7}+5|x_2+2|^{1.9}+12;$
$6|x_1+4|^{0.4}+5|x_2+4|^{1.6}+18;\ 4|x_1+4|^{1.8}+4|x_2+8|^{0.9}+20;$
$6|x_1+8|^{1.6}+5|x_2+8|^{2.4}+22;\ 5|x_1+8|^{1.5}+7|x_2+4|^{0.6}+16;\ 7|x_1+6|^{1.1}+6|x_2+6|^{0.9}+4\}$;
$\mu=4$: $f_4(x_1,x_2)=\min\{4|x_1-2|^{1.5}+5|x_2+2|^{1.9}+11;\ 6|x_1-2|^{2.2}+7|x_2+6|^{2.2}+17;$
$3|x_1-2|^{1.2}+3|x_2+10|^{0.8}+9;\ 7|x_1-6|^{1.7}+7|x_2+10|^{1.7}+19;$
$5|x_1-10|^{0.6}+5|x_2+10|^{1.5}+13;\ 4|x_1-10|^{2.4}+5|x_2+6|^{2.4}+23;$
$6|x_1-10|^{1.1}+7|x_2+2|^{1.6}+15;\ 2|x_1-6|^{1.5}+5|x_2+2|^{1.5}+21;$
$2|x_1-4|^{2.5}+3|x_2+4|^{1.5}+18;\ 3|x_1-4|^{1.3}+2|x_2+8|^{2.3}+10;$
$5|x_1-8|^{1.7}+4|x_2-8|^{0.7}+14;\ 3|x_1-8|^{2.1}+4|x_2+4|^{1.1}+22;\ 5|x_1-6|^{1.1}+5|x_2+6|^{1.3}+6\}$.

Each of them corresponds to value of discrete variable. The required global minimum is in point $\mu^*=1: x^*=(6;\ 6)$, и $f_1(x^*)\equiv f(x^*,y_1^*)=0$.

Each of functions has the inequality restrictions, which allocate the admissible region in circle and on its border:

$\mu=1: (x_1-6)^2+(x_2-6)^2-25\le 0;\ \ \mu=2: (x_1+6)^2+(x_2-6)^2-25\le 0;$
$\mu=3: (x_1+6)^2+(x_2+6)^2-25\le 0;\ \ \mu=4: (x_1-6)^2+(x_2+6)^2-25\le 0.$

For $\mu^* = 1$ spatial form of the minimized function, the lines of equal levels for it, restriction and first steps of movement to global minimum are presented on Figure 1. The conditional global extremum is in point: $\mu^* = 1: x^* = (6;\ 6)$.

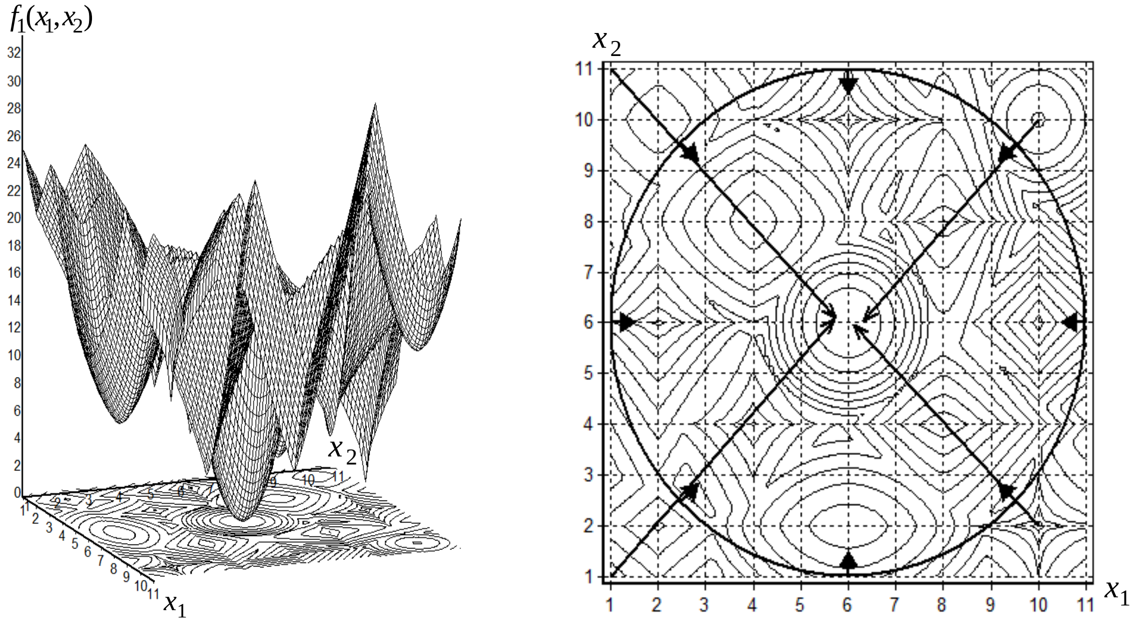

Figure 1. A perspective view of the first function (for $\mu = 1$), equipotential lines, permissible area and first steps of algorithm

Let's show operability of the offered algorithm. The complex characteristic of process of search of global minimum is the estimate $\widehat{P}_{\text{correct}}$ of probability of hitting of the obtained required variables in ε neighborhood of the true solution. The received solution x^* gets in ε neighborhood of true position of global minimum x^{**}, if:

$$\frac{|\, x_\nu^* - x_\nu^{**} \,|}{\Delta x_\nu^0} \le \varepsilon,\ \nu = \overline{1,k}\,, \qquad (7)$$

i.e. if the modules of relative misalignments for all coordinates don't output abroad ε.
Here: Δx_ν^0 – components which determine the sizes of initial rectangular domain of trial movements.

Let's consider dependence $\widehat{P}_{\text{correct}}$ from number of sampling points n. For this purpose N realization of process of search of minimum carried out. The estimate of probability $\widehat{P}_{\text{correct}}$ will be equal to the relative frequency of hitting of the obtained solutions $x_j^*, j = \overline{1,N}$ in rectangular neighborhood of true solution.

The parameters of the algorithm of minimization the following: $x^0 = (0;0)$, $\Delta x^0 = (10;10)$, $\gamma_q = 1$, $q = 2$, kernel on the minimized function parabolic with degree of selectivity $s = 100$. Research parameters: number of realization $N = 101$, size of neighborhood of true solution $\varepsilon = 0.0005$. The received dependence $\widehat{P}_{\text{correct}}$ (by $\mu = 1$) from number of sampling points n is presented on Figure 2.

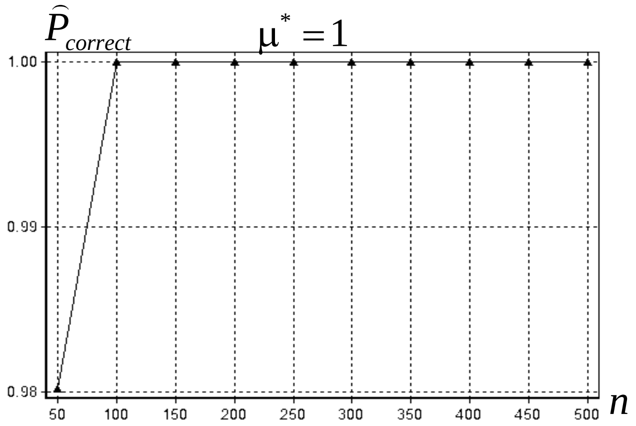

Figure 2. Dependence of estimate of probability of hitting to neighborhood of the true solution from a number of sampling point n

The estimate of probability of hitting of the obtained solution in neighborhood of the true solution is equal to 1 by $n \geq 100$.

5. Statement of problem 2

The problem of search of a conditional global minimum of objective function of many variables $f(x, y)$ with unordered non-numerical possible values at the same inequality restrictions, i.e. at functions of restrictions, which not depend on discrete variable $[\varphi_j(x) \leq 0,\ j = \overline{1, m}]$ is solved:

$$f(x, y) = \min, \quad \varphi_j(x) \leq 0,\ j = \overline{1, m}. \tag{8}$$

We pass from solution of problems of global optimization (3) to solution of one problem:

$$f(x) = \min, \quad \varphi_j(x) \leq 0,\ j = \overline{1, m}. \tag{9}$$

Here in each fixed point x the function

$$f(x) = \min\{ f_1(x), f_2(x), ..., f_r(x)\}. \tag{10}$$

After finding of position x^* of global minimum of function $f(x)$ we define number μ^* of optimal value $y^*_{\mu^*}$ of discrete variable y:

$$\min\{ f_1(x^*), f_2(x^*), ..., f_r(x^*)\} = f_{\mu^*}(x^*). \tag{11}$$

Computing costs at decision reception are reduced almost by r times.

6. Example 2

Let's consider the same problem, as in an example 1 but we set single admissible region inside and on border of circle $x_1^2 + x_2^2 - 121 \leq 0$ radius 11 with the center in origin of coordinates. A perspective view of the minimized function $f(x_1, x_2)$ (look (10)), the lines of equal levels, restriction and first steps of movement to global minimum are presented on Figure 3.

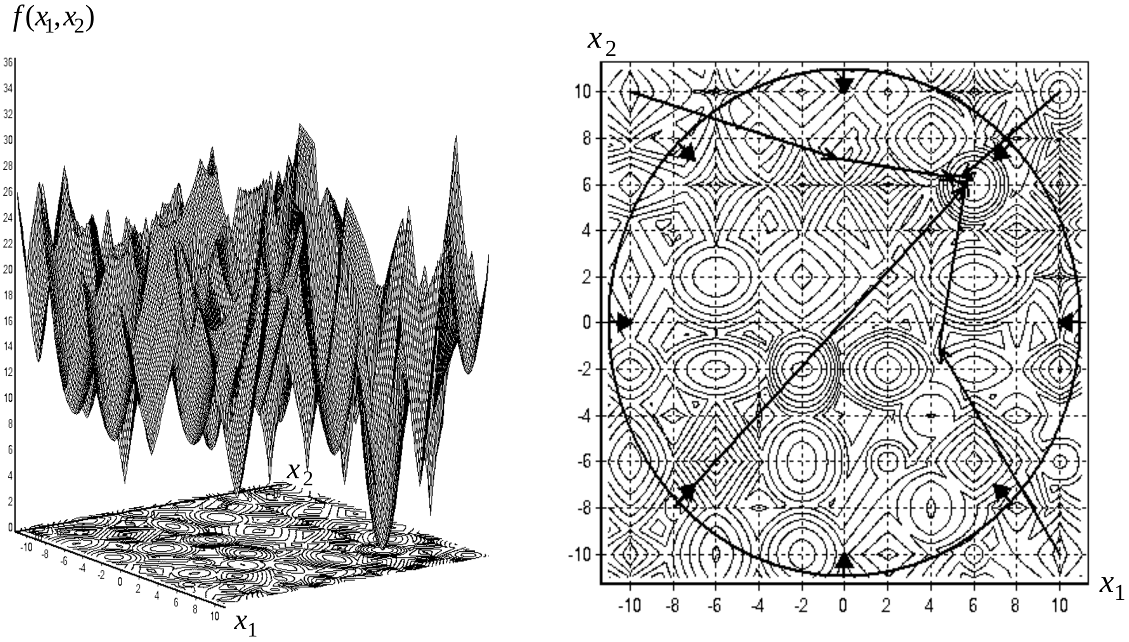

Figure 3. A perspective view of function $f(x_1, x_2)$, equipotential lines, permissible area and first steps of algorithm

Using the scheme, described in section 6, we investigate dependence of estimate $\widehat{P}_{\text{correct}}$ of probability of hitting of the obtained required variables in ε neighborhood of the true solution from a number n of sampling point at $50 \le n \le 500$ and $\varepsilon = 0.01$.

The parameters of the algorithm of minimization the following: $x^0 = (0;\ 0)$, $\Delta x^0 = (10;\ 10)$, $\gamma_q = 1,\ q = 2$, kernel on the minimized function parabolic with degree of selectivity $s = 100$. The number of realization $N = 101$. The received dependence $\widehat{P}_{\text{correct}}$ (by $\mu = 1$) from number of sampling points n is presented on Figure 4.

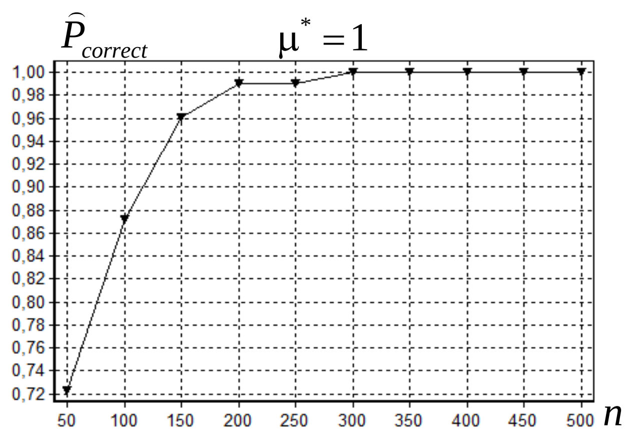

Figure 4. Dependence of estimate of probability of hitting to neighborhood of the optimum true from a number of sampling point n

Apparently from graph on Figure 4 with increase of volume of sampling the estimate of probability of hitting of the obtained solution in neighborhood of the true solution increases and reaches 1 при $n \geq 300$.

7. Conclusion

The offered approach to synthesize algorithms of global optimization at the mixed required variables can be applied to solution: 1) of problems of multicriteria global optimization, 2) of systems of nonlinear equations, 3) of searching of principal minima.

The developed method of global optimization allows to count simply enough on each working step a site of a global extremum. A method kernel is selective averaging of required variables. At the expense of such averaging is carried out more exact movement to an extremum. Transition in algorithms to dimensionless (relative) variables essentially increases speed of convergence, simplifies structure of algorithms and accordingly reduces quantity of a tuned parameters in them. All it is especially important at the account of a considerable quantity of heterogeneous restrictions and at the decision of problems of multicriterion global optimization.

References

[1] Evtushenko Yu G 1971 Numerical methods for finding global extrema (case of a non-uniform mesh) *J. Computational Mathematics and Mathematical Physics* **11:6** 38-64

[2] Strongin R G 1978 *Numerical methods in mulextreme problems* (Moscow: Publishing House «Nauka») p 239

[3] Vasilyev F P 1980 *Numerical methods for solving extreme problems* (Moscow: Publishing House «Nauka») p 520

[4] Polyak B T 1983 *Introduction in optimiztion* (Moscow: Publishing House «Nauka») p 384

[5] Zhiglyavskii A A and Zilinskas A G 1978 *Methods of search for global extremum* (Moscow: Publishing House «Nauka») p 247

[6] Sergeev J D and Kvasov D E 2008 *Difgonal global optimiztion methods* (Moscow: Publishing House «Fizmalit») p 352

[7] Panteleev A V 2009 *Metaheuristic algorithms of search of global extremum* (Moscow: Publishing House «Aviation Institute-Print») p 159

[8] Leyffer S 2001 Integration SQP abd branch-and-bound for mixed integer nonlinear programming *J. Comput. Optim.* **18** 295–309

[9] Fletcher R and Leyffer S 1994 Solving mixed integer programs by outer approximation *J. Math. Program* **66** 327–349

[10] Geoffrion A M 1972 A generalized Benders decomposition *J. of Optimization Theory and Applications* **10(4)** 237–260

[11] Westerlund T and Pettersson T 1995 An extended cutting plane method for solving convex MINLP problems *J. Computers Chem. Eng*. **19** 131–136

[12] Mawengkang H and Murtagh B A 1986 Solving nonlinear integer programs with large-scale optimization software J. *Annals of Operations Research* **5** 425–437

[13] Rouban A I 2003 Method for global optimization in continuous space *J. AMSE Series: Advances A* **40: 4** 9–28

[14] Rouban A I 2004 *Global optimization by a method of averaging of coordinates* (Krasnoyarsk: State Technical University Publisher) p 303

[15] Rouban A I 2013 Global optimization method based on the selective averaging coordinate with restrictions *J. Control and Computer Science* **1(22)** 114-123